Wavefront Reconstruction Methods for a Natural Guide Star Adaptive Optics Application to the Keck Telescope

Mark Milman, Laura Needels, and David Redding

Jet Propulsion Laboratory

California Institute of Technology

1. Introduction. Keck telescope is planning to utilize adaptive optics technology to improve the resolution of the instrument. Telescopes operating in the atmosphere are limited by the seeing conditions at the telescope observational site. An excellent seeing site such as Mauna Kea affords diffraction limited telescope performance to aperture sizes of approximately .3m. The consequence of this limit on astronomical observation is that although a large telescope such as Keck has tremendous light gathering power, the resolution of the 10m instrument is not significantly greater than the resolution of a .3m telescope.

The objective of this study is to investigate and compare slope and curvature wavefront sensing methods for a proposed natural guide star adaptive optics system for the Keck telescope. Specifically we will evaluate the error in wavefront reconstruction resulting from these wavefront sensing mechanisms. A model of the Keck telescope is used to corroborate the analysis and to simulate the curvature sensor; A brief summary of the paper follows.

The second section analyzes the curvature sensing method. The analysis includes nonlinear-, diffraction, and noise effects. Within a geometric model of intensity propagation it is shown that the nonlinearities of the sensor can be characterized through the Gaussian curvature of the wavefront. The scale of this nonlinearity grows with the sensitivity of the measurement. An expression characterizing the balance between the nonlinearity and noise characteristics is derived. It is shown that diffraction effects can be incorporated virr a convolution of the curvature signal with the point spread function of the telescope.

Section 3 deals with reconstruction error for slope sensing and curvature sensing. Covariance matrices of the reconstruction error for both sensing methods are developed. For the case of square arrays analytic expressions are given for the reconstructed wavefront variance. It is shown that these variances are determined from the eigenvalues of the Laplacian operator discretized by a 5- point scheme. Because these eigenvalues are known for the square geometry, the variances can be simply calculated.

In the fourth section a more detailed connection between wavefront reconstruction and correction is developed to analyze the trades between curvature and slope sensing. It is shown that the contribution of sensor/reconstruction error to the correction error depends on the actuator placement. This is in contrast to most analyses where fitting error and sen.sor/reconstruction error are treated independently. The numerical results comparing curvature and Hartmann sensor wavefront reconstruction indicated that for the scale of the Keck adaptive optics system under consideration[1] (approximately 100 subapertures), Hartmann sensor estimation error was superior to curvature sensor estimation error. However, it was observed that curvature sensing compared more favorably as the number of sensors increased as opposed to decreased. This result is somewhat contrary to what has been previously reported[3]. This phenomenon is explained in terms of the *accuracy* (in the sense that the Laplacian is a second order measurement while a gradient is a first order measurement) of the two sensing methods, and an asymptotic expression is given for the estimation variance for both schemes.

2. Curvature Sensing.

The method of curvature sensing using intensity measurements to approximate the Laplacian has been described in several papers[4,5,6]. Briefly, this method relies cm forming a normalized difference of intensities in two planes symmetrically displaced from the focal plane. The derivative of the intensity in the direction of propagation is shown to be proportional to the Laplacian of the wavefront surface. Thus, intensity measurements in displaced planes normal to the direction of propagation provide a finite difference approximation to the differential intensity, and hence to the wavefront Laplacian. The Laplacian together with an estimate of the wavefront slope normal to the boundary is sufficient to pose a standard Neumann problem for estimating the wavefront. In this manner curvature sensing provides an alternative method for wavefront reconstruction. Of particular interest here is the performance of this method as a function of displacement from the focal plane. The trades investigated are between sensor sensitivity and sensor nonlinearity. We will also investigate how diffraction affects the curvature signal. We begin by discussing how the signal gives an approximation to the Laplacian.

Fix a plane P transverse to the direction of propagation. Let the function z represent the wavefront error function, the deviation of the wavefront from being a plane wave propagating in the $z-$ direction. Let $I_0(x, y)$ denote the intensity of the signal at a point $(x, y) \in P$. From geometric optics, the intensity $I_\pm$ at a point displaced a distance $\pm\epsilon$ along the *normal* to the wavefront surface is given by

$$I_\pm = \frac{I_0}{1 \mp 2\epsilon H + \epsilon^2 K}, \tag{1}$$

where H and K are the mean and Gaussian curvatures of the surface, respectively,

$$H_z = \frac{\Delta z}{\sqrt{1 + z_x^2 + z_y^2}}; \quad K_z = \frac{z_{xx} z_{yy} - [z_{xy}]^2}{\sqrt{1 + z_x^2 + z_y^2}}. \tag{2}$$

Here Δz denotes the Laplacian, $\Delta z = z_{xx} + z_{yy}$. Expression (1) can be deduced from the intensity law of geometrical optics' and curvature formulas for parallel displacement of surfaces[8]. Define the normalized difference Q as

$$Q = \frac{I_+ - I_-}{I_+ + I_-}. \tag{3}$$

It is straightforward to show that so long as $1 \mp 2\epsilon H + \epsilon^2 K < 0$,

$$Q = \frac{2\epsilon H}{1 + \epsilon^2 K}. \tag{4}$$

This together with the assumption of paraxial rays leads to the transport equation

$$\frac{\partial I}{\partial z} = \nabla I \cdot \nabla z + I_0 \Delta z. \tag{5}$$

The transport equation above is also valid under paraxial physical optics assumptions[4].

For most applications the gradient of the intensity ∇I vanishes across the aperture and (5) involves only the Laplacian term. The curvature signal is usually modeled by the transport equation, although the expression for Q captures the finite difference approximation (within the geometric optics model for now) that is made in the actual implementation of the sensor. Using the transport equation requires small displacements because it is a linear approximation. The purpose here is to show the trade between displacements, nonlinearities, and sensitivity, and thus (4) is more applicable. The sensitivity of the sensor will be shown to improve for large displacements (big ϵ), however the nonlinearity from the Gaussian curvature term will then become more significant.

Consider an implementation of the curvature sensor where the entrance pupil is imaged in the symmetrically placed planes F_- and F_+, each at a distance l from the focal plane F. Let f denote the system focal length. Then the value of the ϵ term in (4) can be shown to be $\epsilon \approx f^2/l$, for $l << f$. This latter condition is typically easily satisfied. (For the Keck application $f = 150m$ and $l < .1\ m$.) Within the assumptions of geometric optics and paraxial rays the following relationship between the normalized intensity difference in the image planes and the Laplacian at a point (x, y) in the pupil plane holds:

$$\frac{I_+(-lx/f, -ly/f) - I_-(lx/f, ly/f)}{I_+(-lx/f, -ly/f) + I_-(lx/f, ly/f)} = \frac{\epsilon \Delta z}{1 + \epsilon^2 K_z}, \tag{6}$$

where K_z again denotes the Gaussian curvature of the wavefront 2.

Next we will investigate how noise affects this signal. Assume a sensor integration time of ΔT seconds, and let A denote the area of the detector in the planes F_- and F_+. The number of photons captured by these detectors over the period ΔT is modeled as independent Poisson processes N_+ and N_- with means $\bar{N}_+$ and $\bar{N}_-$, respectively. The mean signal intensities in the two planes are then

$$I_\pm = \frac{\bar{N}_\pm}{\Delta T A} \tag{7}$$

Let $I_0 = (I_+ + I_-)/2$, rind define the random variable $\hat{S}$ by

$$\hat{S} = \frac{N_+ - N_-}{2 I_0 \Delta T A}. \tag{8}$$

Then

$$E[\hat{S}] = \frac{I_+ - I_-}{I_+ + I_-}, \tag{9}$$

and the variance of the estimate is computed using the Poisson statistics as

$$E[(\hat{S} - E(\hat{S}))^2] = \frac{1}{N_+ + N_-}. \tag{10}$$

From *(3), (4), and (9)* observe that

$$E(\hat{S}) = \epsilon \Delta z + \phi(z) \qquad \text{with} \quad \phi(z) = \frac{\epsilon^3 \Delta z K_z}{1 + \epsilon^2 K_z} \tag{11}$$

and consequently

$$\Delta z = E(\hat{S}/\epsilon) + \phi(z)/\epsilon. \tag{12}$$

Now define the random variable $\eta = \hat{S}/\epsilon - E(\hat{S}/\epsilon)$. The measured signa from (8) is

$$y = \hat{S}/\epsilon, \tag{13}$$

and thus WC arrive at the sensor model

$$y = \Delta z - \frac{\epsilon^2 \Delta z K_z}{1 + \epsilon^2 K_z} + \eta \tag{14}$$

with

$$E(\eta) = 0 \quad \text{and} \quad E(\eta^2) = \frac{1}{\epsilon^2 (N_+ + N_-)}. \tag{15}$$

Hence the balance that must be maintained is to keep the noise level small by choosing ϵ to be as large as possible, while keeping the nonlinearities at bay with ϵ sufficiently small.

3.3. **Diffraction Effects.** To understand how diffraction modifies the geometric model of curvature sensing, let, $h_+(x)$ denote the impulse response between the object plane 0 and the image plane F_+. Diffraction effects at the plane F_+ can be modeled by convolving the intensity $I_+(x, y)$ with the squared modulus of the impulse response[9],

$$I_+^{diff}(x_+, y_+) = \int_{\Sigma_+} |h_+(x_+ - x, y_+ - y)|^2 I_+(x, y)\, dx\, dy, \tag{16}$$

where $I_+(x, y)$ is the geometric optics prediction] of the intensity in the image plane and Σ_+ is the *support* of I_+ (where it is nonzero). For $1 << f$ wc can replace h_+ with h, where h denotes the impulse response in the focal plane. Diffraction effects in the plane F_- can be approximated in a similar manner:

$$I_-^{diff}(x_-, y_-) = \int_{\Sigma_-} |h(x_- - x, y_- - y)|^2 I_-(x, y)\, dx\, dy, \tag{17}$$

where $I_-(x, y)$ is the geometric optics intensity ill F_-,

Let χ denote the characteristic function of a set ($\chi(S)(x) = 1$ if $x \in S$, O otherwise), and define the normalized impulse response $\tilde{h}$

$$|\tilde{h}(x, y)|^2 = \frac{|h(x, y)|^2}{\int |\tilde{h}(x, y)|^2 dx\, dy}.$$

The following expression for the diffracted curvature signal is obtained

$$\frac{I_{+}^{diff}(x,y) - I_{-}^{diff}(x,y)}{I_{+}^{diff}(x,y) + I_{-}^{diff}(x,y)} \approx \frac{\epsilon\tilde{h} * \Delta z[(1 - 2\epsilon^2 K)\chi(\Sigma_+ \cap \Sigma_-)] + 1/2\tilde{h} * \chi(\Sigma_+ - \Sigma_-) - 1/2\tilde{h} * \chi(\Sigma_- - \Sigma_+)}{\tilde{h} * (1 - \epsilon^2 K)\chi(\Sigma_+ \cap \Sigma_-) + 1/2\tilde{h} * \chi[(\Sigma_+ - \Sigma_-) \cup (\Sigma_- - \Sigma_+)]}$$

where * denotes convolution. Thus for points far removed from the pupil edges or obscurations (i.e, well within the interior of $\Sigma_+ \cap \Sigma_-$), the curvature signal is characterized hy convolving the individual terms of the geometric model with the point spread function of the instrument. For systems with large aperture, $\tilde{h}$ is an approximate δ function, and we recover the geometric model.

Modifying the sensor model to include diffraction effects is straightforward. For points in the interior of $\Sigma_+ \cap \Sigma_-$, the estimator is not estimating the Laplacian of the wavefront, Δz, hut the convolution of the the Laplacian with the normalized point spread function of the instrument. Thus the model becomes

$$y = \tilde{h} * \Delta z + \tilde{h} * \phi + \eta, \tag{20}$$

where * again denotes convolution, ϕ is the nonlinear term from (14) and the noise term η has the same statistics as before. Closer to the boundaries of the obscurations and pupil edges the quotient model (19) must be used. (The quotient model also explains how the intensity signal can be used for estimating the radial derivatives on the boundary.)

Examples of the Curvature Signal. The Zernike polynomial for tilt is

$$z(x,\ y) = \frac{x}{R} \quad \text{(or } z = \frac{\rho}{R} cos\theta \text{ in polar coordinates);} \quad \text{R= pupil radius.}$$

Hence,

$$\Delta z = 0.$$

The zero Laplacian of tilt is captured by the intensity signal in Figures 1a- 1c. In each of these figures we chose the displacement 1 from focus to be .05m, and the focal length of the system, f= 150m. Thus $\epsilon = 4.5 \times 10^5$. Figure 1 a contains the Keck prescription without central obscuration from the secondary mirror. The magnitude of the signal increases to unity at the edge of the pupil. From geometric optics considerations, the width of the signal where it approximates unity can be shown to be proportional to the radial derivative of the wavefront aberration. Figures 1b 1c contain the signal with obscuration. There is more ringing to the signal in these cases because of the di fir-action contribution of the secondary mirror. The signal in the center for Figures 1b- 1c is due entirely to diffraction. Because the aberration consists of an x-- axis tilt, the terms containing the Laplacian and Gaussian curvature in (19) disappear. Along the y- axis we would expect the signal to diminish in the obscured region because the terms in the numerator cancel. This is precisely the case as can be oh.served in Figure 1c.

The Zernike polynomial for defocus is given by

$$z(x,\ y) = 2(\frac{x}{R})^2 + 2(\frac{y}{R})^2 - 1; \quad \text{R= pupil radius}$$

Thus with a coefficient of ρ multiplying the defocus term, the resulting Laplacian is

$$\epsilon \Delta z = \frac{8\rho}{R^2}.$$

The corresponding geometric curvature signal for defocus is given hy

$$\frac{I_+(x,y) - I_-(x,y)}{I_+(x,y) + I_-(x,y)} = \frac{-\frac{f^2}{l}[\frac{8\rho}{R^2}]}{1 + (\frac{f^2}{l})^2 \frac{16\rho^2}{R^4}}.$$

In both cases the signal is constant. The nonlinearity introduced by the Gaussian curvature term in the denominator is seen in the simulations by comparing intensities from normalized intensity maps as the image planes are moved closer to focus (Figs. 2a- 2c). The intensities increase sublinearly because of this term. In

Table 1 the values predicted by linear theory (5), the geometric signal (4) and the simulations are compared. The geometric prediction and the simulations agree very well, while the linear estimate overestimates the signal as the image planes are moved closer to focus.

4. **Wavefront Reconstruction.** The problem of wavefront reconstruction is to estimate the wavefront across an aperture from sampled values. The sample values are obtained from measurement devices such as a Hartmann sensor, a shearing interferometer, or curvature sensor. These sensors do not provide direct information of the wavefront, but only of first or second derivative information through either slope or curvature measurement.

The general setup of the reconstruction problem is fairly simple. Let an aperture be defined by a region D in a plane with boundary ∂D. The reconstruction problem for slope measurements is to determine $W(x), x \in D$ given a sample of the gradient of W, $\nabla W(x_i), i = 1, \ldots, n$. The problem for curvature sensing is to estimate $W(x)$ given the samples $\Delta W(x_i), i = 1, \ldots, n$. For simplicity we will assume the region in the plane is a square, (It can be shown that with little penalty in estimation error, the reconstruction problem can always be imbedded into a square, from where it is possible to exploit the specific structure of the estimation problem[10].)

We will assume that the square is $d \times d$ units and there are $(N+1)^2$ regularly spaced nodes. We let h denote the mesh width, so that $h = d/N$. We begin with an analysis of slope measurements in this configuration. At each mesh point consider the noisy vector of slope measurements

$$[s^x_{ij} \quad s^y_{ij}] = \nabla W(x_i, y_j) + \eta_{ij} \tag{21}$$

where

$$\eta_{ij} = [\eta^x_{ij} \quad \eta^y_{ij}], \tag{22}$$

with η_{ij} zero mean for every i and j, and with constant, covariance $E(\eta^T_{ij}\eta_{ij}) = \sigma^2 I_{2\times 2}$. (Under reasonable assumptions the x and y slope measurements can be treated as independent[11].)

The gradient $\nabla W(x_i, y_j)$ is approximated at interior mesh points by the difference operator A_h,

$$(A_h u)(ij) = \left(\frac{u_{ij+1} - u_{ij}}{h}, \frac{u_{i+1j} - u_{ij}}{h}\right) \tag{23}$$

To develop the minimum variance estimator we write the difference operator above as

$$A_h = \frac{1}{h}\begin{bmatrix} A^x \\ A^y \end{bmatrix} \tag{24}$$

where

$$(A^x u)(ij) = u_{ij+1} - u_{ij},$$

and

$$(A^y u)(ij) = u_{i+1j} - u_{ij}.$$

If A has independent columns, the minimum variance solution is obtained as

$$\hat{u} = (A^T A)^{-1} A^T y. \tag{25}$$

Observe that since A has a nontrivial kernel spanned by the single vector $v = [1 \ldots 1]^T$, $A^T A$ is not invertible. This is merely a uniqueness problem, and one way of fixing solutions is to define a map $\Psi : R^{(N+1)^2-1} \to R^{(N+1)^2}$ so that the range of Ψ is the orthogonal complement, call it U, of the subspace spanned by v. We can now repose the problem as finding the linear minimum variance estimate in the subspace U:

$$\min_{\hat{u}} E(|\hat{u} - u|^2); \quad y = A_h u + \eta \quad \text{where} \quad u \in U. \tag{26}$$

This is equivalent to the problem

$$\min_{\hat{w}} E(|\Psi(\hat{w} - w)|^2); \; y = A_h \Psi w + \eta \tag{27}$$

This reposed problem has the interpretation that we are seeking solutions to the problem where the "mean" wavefront is zero. We could have chosen any other matrix, say Γ: $R^{(N+1)^2-1} \to R^{(N+1)^2}$, such that $A\Gamma$ has full rank to make the problem well posed. However, it can be shown that the choice Ψ leads in a certain sense to the minimum variance *wavefront error* solution. Writing $\hat{u}_\Psi$ and $\hat{u}_\Gamma$ for the solutions in the subspaces, $R(\Psi)$ and $R(\Gamma)$, respectively, it can be shown that

$$E(|\hat{u}_\Psi - u_\Psi|^2) = E\{|\hat{u}_\Gamma - u_\Gamma|^2 - (< \tilde{v}, \hat{u}_\Psi - u_\Psi >)^2\},$$

where $\tilde{v} = v/|v|$, with $v = [1 \ . \ . \ . \ 1]^T$.

Choosing Ψ also facilitates the expression for the variance of the estimate as indicated in the theorem below.

Theorem 1. Let $Q = E(\eta\eta^T)$ denote the measurement noise covariance matrix from (27) and let A_h be the difference operator defined in (24). Then the minimum variance solution is

$$\hat{u} = \Psi\hat{w}, \quad \text{where} \quad \hat{w} = [\Psi^T A_h^T Q^{-1} A_h \Psi]^{-1} \Psi^T Q^{-1} A_h^T y,$$

with variance

$$E(|u - \hat{u}|^2) = tr\{[\Psi^t A_h^T Q^{-1} A_h \Psi]^{-1}\}$$

If Q is a scalar matrix, i.e., $Q = \sigma^2 I$ then the variance can be expressed as

$$E(|u - \hat{u}|^2) = \sigma^2 h^2 tr\{[\Psi^t A^T A \Psi]^{-1}\}$$
$$= \sigma^2 h^2 \sum \frac{1}{\lambda_i},$$

where the $\lambda_i's$ are the nonzero eigenvalues of $A^T A$

We will next see how a very analogous situation develops for curvature sensing when using a 5 point discretization scheme for the Laplacian. A point to keep in mind while we develop the result below is that it is tied to this particular approximation of the Laplacian, and other options for discretization are available for curvature sensing.

Curvature sensing produces the following sampled Laplacian signal:

$$\Delta w_{ij} + \eta_{ij} = y_{ij} \tag{28}$$

Discretizing the Laplacian via the 5-point scheme[12] leads to the difference equation

$$\frac{4w_{ij} - w_{ij-1} - w_{ij+1} - w_{i-1j} - w_{i+1j}}{h^2} = y_{ij} + \eta_{ij}.$$

From this discretization it can be shown that

$$A_h^T A_h w + \eta = y$$

with A_h defined as in (24). Analagous to Theorem 1, we have for curvature sensing reconstruction

Theorem 2. Let Q denote the covariance of the noise tern in (28) and let A_h be the difference operator defined in (24). Then the minimum variance solution is

$$\hat{u} = \Psi(\psi^T A_h^T A_h \Psi)^{-1} \Psi^T y,$$

with variance

$$E(|u - \hat{u}|^2) = tr\{[\Psi^T A_h^T A_h Q A_h^T A_h \Psi]^{-1}\}$$

If Q is a scalar matrix, i.e., $Q = \sigma^2 I$ then the variance can be expressed as

$$E(|u - \hat{u}|^2) = \sigma^2 h^4 tr\{[\Psi^T A^T A A^T A \Psi]^{-1}\}$$
$$\sigma^2 h^4 \sum \frac{1}{\lambda_i^2},$$

where the $\lambda_i's$ are the nonzero eigenvalues of A^TA.

There are two differences between the variances for slope and curvature sensing. The first is the factor of h^4 that appears in the curvature sensor reconstruction error, versus the factor of h^2 in the slope sensing reconstruction error. The second difference is that the reconstruction in curvature reconstruction involves the sum over the square of the reciprocals of the eigenvalues of A^TA, ***as*** opposed to the reciprocals of the eigenvalues for slope sensing. Thus we see immediately that the trade between curvature and slope sensing is governed by both the growth of the reciprocals of the eigenvalues of the Laplacian ***and*** the mesh size.

On square domains it is fortunate that these eigenvalues can be computed analytically so that direct comparisons between the two methods can be made.

Proposition 3. The eigenvalues of A^TA are given by

$$\lambda_{ij} = 4 - 2cos\frac{\pi i}{N+1} - 2cos\frac{\pi j}{N+1}, \quad i,j = 0,\ldots,N.$$

This result will be called upon in the next section

4, **Estimation Error, Curvature vs. Hartmann Sensor Comparison.** In this section we will investigate and compare the estimation error for Hartmann and curvature sensors. Let $w(x)$ denote the instantaneous wavefront and $u(x)$ the corrected wavefront surface. It will be assumed that $u(x)$ is formed from the actuator response to the *reconstructed* wavefront based on the sensor information, (This will be made precise below) The mean square error is

$$J = \frac{1}{A(\Delta)} \int_\Delta |w(x) - u(x)|^2 dx, \tag{29}$$

where Δ is the aperture of the deformable mirror, and $A(\Delta)$ denotes its area. ($2\pi\sqrt{J}/\lambda$ gives the rms wavefront error in radians of phase, where λ denotes the sensing wavelength wavelength.) Now $u(x)$ is developed as a linear function of the estimated sampled wavefront vector $\hat{w}$ obtained via the reconstruction process. Thus we write $\hat{w} = [\hat{w}(x_1 \ \hat{w}(x_N)]^T$, where $\hat{w}(x_i)$ is an estimate of the wavefront at x_i. For each wavefront function $w(x)$, let δw denote the vector $[w(x_1)\ldots w(x_N)]^T$. Let $L_2(\Delta)$ denote the space of square integrable real valued functions on Δ with inner product

$$< f,g > = \int_\Delta f(x)g(x)dx.$$

Since u is a linear function of $\hat{w}$ there exists a linear operator $T : R^N \to L_2(\Delta)$ such that

$$u = T\hat{w}.$$

If there was no sensor or reconstruction error we would have $\hat{w} = \delta w$, and J would simply reduce to the fitting error J_{fit}

$$J_{fit} = \frac{1}{A(\Delta)} \int_\Delta |w(x) - T\delta w|^2 dx. \tag{30}$$

Next let E denote the expectation operator and observe that

$$E(J) = \frac{1}{A(\Delta)}\{\int_\Delta |w(x) - T\delta w|^2 dx + \int_\Delta E[|T(\delta w - \hat{w})|^2]dx +$$
$$\int_\Delta E\{(w - T\delta w)(T(\delta w - \hat{w}))\}dx\}$$

The first integral on the right above is recognized as the fitting error J_{fit}. Note that the last integral vanishes since $\hat{w}$ is an unbiased estimate of δw, i.e., $E(\delta w - \hat{w}) = 0$. Thus,

$$E(J) = J_{fit} + \frac{1}{A(\Delta)} \int_\Delta E[|T(\delta w - \hat{w})|^2]dx. \tag{31}$$

So now we will examine the integral term above. Observe that since $T : R^N \to L_2(\Delta)$, there exist "influence" functions $\{\Psi_i\}_{i=1}^N \subset L_2(\Delta)$ such that $Te_i = \Psi_i$, where e_i denotes the vector in R^N with a one in the i^{th} entry and zeros elsewhere. Let T^* denote the adjoint of T,

$$< T^* v, \alpha > = < v, T\alpha >, \quad \text{for all } v \in L_2(\Delta), \alpha \in R^N,$$

and note that $T^*T : R^N \to R^N$ is represented as a matrix with ij^{th} entry $(T^*T)_{ij}$

$$(T^*T)_{ij} = \int_{\Delta} \Psi_i \Psi_j dx. \tag{32}$$

Hence,

$$\frac{1}{A(\Delta)} E \int_{\Delta} |T(\delta w - \hat{w})|^2 dx = \frac{1}{A(\Delta)} E < T^*T(\delta w - \hat{w}), (\delta w - \hat{w}) >_{R^N} . \tag{33}$$

Now let Σ denote the $N \times N$ covariance matrix of the estimate of δw,

$$\Sigma = E[(\delta w - \hat{w})(\delta w - \hat{w})^T]. \tag{34}$$

Thus we obtain

$$\frac{1}{A(\Delta)} E \int_{\Delta} |T(\delta w - \hat{w})|^2 dx = \frac{1}{A(\Delta)} tr[\Sigma T^* T], \tag{35}$$

so that

$$E(J) = J_{fit} + \frac{1}{A(\Delta)} tr[\Sigma T^* T]. \tag{36}$$

As with the standard adaptive optics error analysis, the static error contains the sum of the fitting error plus a sensor/reconstruction error term containing the estimation covariance matrix Σ. However, in this analysis it is clear that this sensor/reconstruction error is linked to the spatial correlation matrix of the actuation scheme via the matrix $T^* T$. Therefore, when comparing various sensor/reconstruction methods, the precise term of interest is the estimation error, which we denote as J_{est},

$$J_{est} = \frac{1}{A(\Delta)} tr[\Sigma T^* T]. \tag{37}$$

Although the correlation matrix $T^* T$ can be derived from either analytical models or experimental data, for the purpose of the analysis of this section we make the simplifying assumption that, the influence functions are orthogonal with respect to the L_2 inner product. Thus we use the approximation

$$J_{est} \approx \frac{A(\Delta_i)}{A(\Delta)} tr(\Sigma) \tag{38}$$

For segmented mirrors this assumption is valid. For continuous face sheet mirrors a more realistic and conservative assumption is that $T^* T$ is diagonally dominant. A multiplicative factor slightly greater than unity **would** then appear on the right side above.

5.1. Curvature vs Slope Sensing Error Estimates. Putting the various pieces of the analysis together we can now give some simple estimates for the estimation error J_{est} for slope and curvature sensing.

In the case of a square aperture equipped with an $N \times N$ array of subapertures, the covariance matrix Σ for either curvature or Hartmann sensing is related to the eigenvalues $\{\lambda_{kl}\}_{kl}$ of $A^T A$ given in Proposition 4.3. To apply the formula for J_{est} it is necessary to compute the variances

$$tr(\Sigma_{slope}) = \sigma^2_{slope} \sum_{k,l} \frac{1}{\lambda_{kl}} \quad \text{and} \quad tr(\Sigma_{curve}) = \sigma^2_{curve} \sum_{k,l} \frac{1}{\lambda^2_{kl}},$$

where σ^2_{slope} and σ^2_{curve} are the variances for slope and curvature sensing, respectively. Assuming a circular subaperture,

$$\sigma_{slope} = 3\pi\lambda I(h,\lambda)/16h\sqrt{N},$$

where λ is the operating wavelength, h = subaperture diameter, N is the number of photons per subaperture, and $I(\lambda, h)$ is a seeing related term which gives the slope variance correction factor as a function of h/r_0. $I(h, \lambda)$ is near unity for $h < 70$[13]. We will take the previous values, $f = 150m$, and $1 = .05m$ to compute σ_{curve}. Reconstruction errors were compared for various array sizes with N = 5, 10, 20, and 40 corresponding Lo mesh widths of h = 2m, 1m, .5m, and .25 m, respectively. These results are contained in Table 2.

Table 2. Curvature vs. Slope Sensing Reconstruction

h	.25	.5	1.0	2,0
Hartmann	***5.84*** x 10- ⁷	***7.87*** x 10-7	1.23x 10- ⁶	2.13x 10- ⁶
Curvature	9.0'7 x 10'7	1.82 X10-6	3.68 X1 O- ⁶	7.60 x 10-6

What is observed in the table is that the curvature sensor compares more favorably with the Hartmann sensor as the resolution *increases*, which is somewhat contrary to what is typically reported[2]. The trade that occurs between the two sensing methods is that although reconstruction error from the Laplacian measurement grows more rapidly than the reconstruction error from gradient measurements, this effect is mitigated by the properly that the Laplacian is a higher accuracy measurement ($O(h^2)$ versus $O(h)$.) This trade is made more evident in the asymptotic estimates below.

For large N the sums in Theorems 1 and **2** can be approximated as

$$\sum_{l,k} \frac{1}{\lambda_{kl}} \approx O(N^2 log(N)),$$

and

$$\sum_{l,k} \frac{1}{\lambda_{kl}^2} \approx O(N^4).$$

From these approximations we obtain asymptotic estimates of the surface reconstruction error for slope sensing and curvature sensing:

Theorem 4. Let d = length of a square aperture, and let h denote the mesh size. Then for large ***d/h,*** the following asymptotic reconstruction error estimates are obtained for slope sensing and curvature sensing, respectively:

$$J_{est}^{slope} \approx \sigma_{slope}^2 h^2 log(d/h), \tag{39}$$

and

$$J_{est}^{curve} \approx \sigma_{curve}^2 h^2 d^2. \tag{40}$$

Most previous comparisons of slope and curvature sensing reconstruction fixes the mesh size h and varies the aperture size ***d***. As ***d*** increases (with h fixed) it is seen that the rms error from slope measurement reconstruction (RMS_{slope}) grows logarithmically and the rms reconstruction error from curvature sensing (RMS_{curve}) grows linearly as reported[1]. The error propogates differently, however, if we fix the aperture size and decrease the mesh size. Recalling that

$$\sigma_{slope} = \frac{3\pi}{16h} \frac{\lambda}{} \frac{1}{\sqrt{N}}, \tag{41}$$

and

$$\sigma_{curve} = \frac{l}{f^2} \frac{1}{\sqrt{N}}, \tag{42}$$

where N denotes the number of collected photons, we find that

$$RMS_{slope} = \frac{3\pi\lambda}{16} \sqrt{log(d/h)} \frac{1}{\sqrt{N}}, \tag{43}$$

and

$$RMS_{curve} = \frac{lhd}{f^2\sqrt{N}}. \quad (44)$$

Thus we see that curvature sensing may actually be advantageous to Hartmann sensing when the subaperture diameters must be small. (Such a circumstance is envisioned for the dense segmented primary mirror of the SELENE telescope[14].)

Acknowledgement

This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

[1] D. Redding, M. Milman, and L. Needels, Large telescope natural guide star adaptive optics system, SPIE Conf. on Astronomical Telescopes and Instrumentation for the 21st Century, Kona, HA, 1994.

[2] C. Roddier, E. Limburg, N. Roddier, F. Roddier, M. Northcott, Interferometric imaging through aberrated optics without a reference source, Annual Report SDI/IST Contract, 1989.

[3] F. Roddier, M. Northcott, and J. E. Graves, "A simple low- order adaptive optics system for near infrared applications", Pub. Astronomical Sot. of Pacific, 103, January, 1991.

[4] M. R. Teague, Deterministic phase retrieval: A Green's function solution, J. Opt. Sot. Am., 73, 1983, pp. 14341441.

[5] F. Roddier, A new curvature method, Applied Optics, 27, 1988, pp. 12231 '225.

[6] N. Streibel, Phase imaging by the transport equation and intensity, Opt. Comm , 49, 1984, pp. 6- 10.

[7] M. Born and E. Wolf, "Principles of Optics", Pergamon Press, Oxford, 1989.

[8] B. O'Neill, "Elementary Differential Geometry", Academic Press, N. Y., 1966.

[9] J. W. Goodman, "Introduction to Fourier Optics", McGraw- hill, N. Y., 1968.

[10] M. Milman, A. Fijany, and D. Redding, Wavefront control analysis for a dense adaptive optic system, SPIE Int. Symposium on Laser Power Beaming, Los Angeles, 1994.

[11] R. H. Hudgin, Wavefront reconstruction for compensated imaging, J. Opt. Sot. Am., 67, 1973, pp. 375 378.

[12] G. Strang, "Introduction to Applied Mathematics", Wellesl - Cambridge Press, Wellesley, MA, 1986.

[13] H. T. Yura, Short- term average optical beam spread in a turbulent medium, J. Opt. Sot. Am., 63, 1973, pp. 567-572.

[14] R. Ulich and J. D. G. Rather, Innovative approach to next generation telescope design, SPIE Conf. 1236, Tuscon, AZ, 1990.